ISSN 0001-4346, Mathematical Notes, 2013, Vol. 93, No. 4, pp. 633–635. © Pleiades Publishing, Ltd., 2013.
Original Russian Text © I. A. Mikhailova, 2013, published in Matematicheskie Zametki, 2013, Vol. 93, No. 4, pp. 634–636.

SHORT COMMUNICATIONS

On the Avoidability Index of Palindromes

I. A. Mikhailova*

Ural Federal University, Yekaterinburg

Received October 10, 2011

DOI: 10.1134/S0001434613030334

Keywords: *avoidable word, morphism, avoidability index, palindrome, fusion, free deletion.*

Let Σ and Δ be two nonempty finite (not necessarily different) sets (*alphabets*). As usual, by Σ^+ (by Δ^+) we denote the set of all nonempty words over the alphabet Σ (respectively, Δ). A map $h\colon \Delta^+ \mapsto \Sigma^+$ is called a *morphism* if $h(p)h(q) = h(pq)$ for any words $p, q \in \Delta^+$. We say that a word $u \in \Sigma^+$ *avoids* a word $p \in \Delta^+$ if, for any morphism $h\colon \Delta^+ \mapsto \Sigma^+$, the word $h(p)$ is not a subword of u. A word p is said to be k-*avoidable* if there exists an infinite sequence of words u_i over some k-letter alphabet such that every word u_i avoids p. Finally, a word p is called *avoidable* if it k-avoidable for some number k; otherwise, the word p is said to be *unavoidable*. The least number k for which a word is k-avoidable is called the *avoidability index* of this word. The problem of finding the avoidability index of an arbitrary avoidable word has not been solved; even the question on the complexity of this problem remains open (see the survey [1]). However, for words of certain types, various bounds for avoidability index have been obtained. For example, Petrov [2] showed that the avoidability index of any complete word is at most 4 (a word p is said to be *complete* if every letter of p occurs twice and, whenever p contains different letters x and y, it also contains xy and yx as subwords). In this note, we specify yet another natural class of words for which the avoidability index has an absolute (i.e., not depending on the number of letters) bound.

Recall that a word $p = a_1a_2\dots a_n$ is called a *palindrome* if it coincides with its mirror image $\overleftarrow{p} = a_n \dots a_2a_1$. We prove the following theorem.

Theorem 1. *The avoidability index of any avoidable palindrome does not depend on the number of letters in this word and is at most* 16.

The avoidable palindrome $abacbdbcaba$ does not belong to any of the previously considered classes of avoidable words, and nothing has been known about its avoidability index so far.

To prove the theorem, we need additional information about avoidable words.

1. FREE DELETIONS

Let u be a word over an alphabet Σ. A pair of subsets $B, C \subseteq \Sigma$ is called a *fusion* in u if we have $x \in B \Leftrightarrow y \in C$ for any two-letter subword xy of u. A set $A \subseteq B \setminus C$ is said to be *free* in the word u. We refer to the removal from the word u of all letters belonging to a free set A as a *free deletion*; we denote this operation by δ_A and the result of applying it to u by u_A. A sequence $\delta_{A_1}, \delta_{A_2}, \dots, \delta_{A_k}$ is called a *sequence of free deletions* if δ_{A_1} is a free deletion in u, δ_{A_2} is a free deletion in u_{A_1}, and so on.

The following theorem from [3] describes a relationship between avoidable words and free deletions in words.

Theorem 2. *A word u is unavoidable if and only if there exists a sequence of free deletions such that the application of this sequence to the word u yields the empty word.*

*E-mail: inna.mikhaylova@gmail.com

Take words $u \in \Sigma^+$ and $p \in \Delta^+$ and a morphism $h\colon \Delta^+ \mapsto \Sigma^+$ such that $h(p)$ is a subword of u. Choose a set $D \subseteq \Sigma$ and let $\overline{D}$ denote $\{x \in \Delta \mid h(x) \in D^+\}$. Consider the new morphism

$$h_D\colon (\Delta \setminus \overline{D})^+ \mapsto (\Sigma \setminus D)^+$$

defined by $h_D(x) = (h(x))_D$. The following lemma was proved in [4].

Lemma 1. *Let $u = u_1 h(p) u_2 \in \Sigma^+$ for a nondeleting morphism h, and let $p \in \Delta^+$. If a set $D \subseteq \Sigma$ is free in the word u, then $\overline{D}$ is free in p and*

$$h_D(p_{\overline{D}}) = (h(p))_D.$$

2. PROOF OF THEOREM 1

The construction of a sequence of words over an alphabet of 16 letters described below is a special case of a construction proposed by Sapir in [4]. Consider the 16×4 matrices

$$P = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 3 & 1 & 3 & 1 \\ 4 & 1 & 4 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 4 & 1 & 4 \\ 2 & 4 & 2 & 4 \\ 3 & 4 & 3 & 4 \\ 4 & 4 & 4 & 4 \end{pmatrix}, \qquad P_\Sigma = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{12} & a_{23} & a_{14} \\ a_{31} & a_{12} & a_{33} & a_{14} \\ a_{41} & a_{12} & a_{43} & a_{14} \\ \vdots & \vdots & \vdots & \vdots \\ a_{11} & a_{42} & a_{13} & a_{44} \\ a_{21} & a_{42} & a_{23} & a_{44} \\ a_{31} & a_{42} & a_{33} & a_{44} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}.$$

We denote the set of letters $\{a_{ij} \mid 1 \le i, j \le 4\}$ by Σ. The matrix P_Σ is obtained by replacing the elements of the matrix P by letters from Σ according to the following rule: if $i, j \le 4$ and the number i occurs in the jth column of the matrix P, then the corresponding element of P_Σ is a_{ij}. For each $1 \le j \le 4$, let us denote the set $\{a_{ij} \mid 1 \le i \le 4\}$ by Σ_j; then $\Sigma = \bigcup_{j=1}^{4} \Sigma_j$.

We denote the word in the ith row of the matrix P_Σ by B_i. The words B_i with $1 \le i \le 16$ are called *blocks* and have the following obvious properties [4].

Lemma 2. *In each block, all letters are different, and for each j, the jth position in this block is occupied by a letter from Σ_j. Two different blocks contain no equal two-letter subwords.*

Consider a morphism $\gamma\colon \Sigma^+ \mapsto \Sigma^+$ such that $\gamma(a_{ij}) = B_{(i-1)4+j}$. Let us show that each of the words $\gamma^m(a_{11})$, where $m \ge 0$, avoids all avoidable palindromes.

Suppose that, on the contrary, there exists a (minimal) number m for which there exists an avoidable palindrome $p \in \Delta^+$ and a morphism $h\colon \Delta^+ \mapsto \Sigma^+$ such that the word $h(p)$ is a subword of $\gamma^m(a_{11})$. Without loss of generality, we can assume that $h(p) = \gamma^m(a_{11})$. Let xy be a two-letter subword of p. Since $p = \overleftarrow{p}$, it follows that yx is a subword of p as well. By virtue of Lemma 2, the second indices of all letters in each block increase from 1 to 4; therefore, $h(x)$ and $h(y)$ can be arranged in the word $\gamma^m(a_{11})$ only as shown in the figure (see below).

Thus, the number of "boundaries" of the images of letters inside a block is at most two, so that each block has a two-letter subword which is either contained in or disjoint from the image of any letter. We fix a set of such two-letter subwords and refer to them as *markers*. By Lemma 2, each two-letter subword uniquely determines the block. To each marker we assign a new letter t_B and denote the set of letter assigned to markers by T.

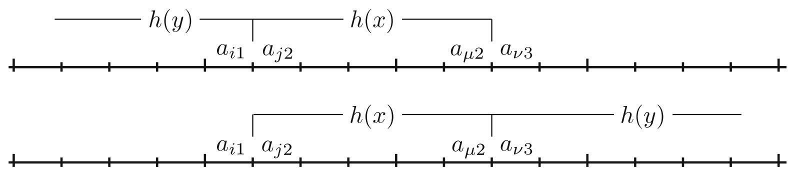

Figure: An illustration to the proof of Theorem 1.

Let $\overline{h}(x)$ denote the word $h(x)$ in which all markers are replaced by the corresponding letters from T. This defines a morphism

$$\overline{h}\colon \Delta^+ \mapsto (\Sigma \cup T)^+.$$

Consider the word

$$\overline{h}(p) = P_1 t_1 Q_1 P_2 t_2 Q_2 \dots P_\ell t_\ell Q_\ell,$$

where P_i (Q_i) is the prefix (suffix) of the corresponding block. Following [4], we say that this word is *quasi-integral* (we use this term because, in [4], those words from Σ^+ which can be divided into blocks were called integral). It was shown in [4] that the sequence of deletions $\delta_{\Sigma_1}, \delta_{\Sigma_2}, \delta_{\Sigma_3}, \delta_{\Sigma_4}$ is free in any quasi-integral word. Applying this sequence to the word $\overline{h}(p)$, we obtain $\Lambda = t_1 t_2 \dots t_\ell \in T^+$.

By Lemma 1, the sequence $\delta_{\Sigma_1}, \delta_{\Sigma_2}, \delta_{\Sigma_3}, \delta_{\Sigma_4}$ generates a sequence of free deletions in the word p; we denote the result of the application of this sequence by p'. Note that, deleting a set of letters in a palindrome, we again obtain a palindrome. Consider the morphism $\varphi\colon T^+ \mapsto \Sigma^+$ defined by the rule

$$\varphi(t_B) = a_{ij} \qquad \text{if} \quad \gamma(a_{ij}) = B.$$

It is easy to see that

$$\varphi(t_1 t_2 \dots t_\ell) = \gamma^{m-1}(a_{11}).$$

By Lemma 1, the image of the palindrome p' under some morphism is a subword of Λ; therefore, the word $\varphi(\Lambda) = \gamma^{m-1}(a_{11})$ does not avoid p'. Recall that the number m was chosen so that $\gamma^{m-1}(a_{11})$ avoids all avoidable palindromes. It follows that the word p' is unavoidable, and, by Theorem 2, the word p is unavoidable as well; this leads to a contradiction and completes the proof of the theorem.

REFERENCES

1. J. D. Currie, Theoret. Comp. Sci. **339** (1), 7 (2005).
2. A. N. Petrov, Mat. Zametki **44** (4), 517 (1988) [Math. Notes **44** (4), 764 (1988)].
3. D. R. Bean, A. Ehrenfeucht, and G. F. McNulty, Pacific J. Math. **85** (2), 261 (1979).
4. M. V. Sapir, *Combinatorics on Words with Applications*, LITP Report 32 (Univ. Paris 7, Paris, 1995).